Virtualization Technologies for Hadoop-based applications

Aymen Jlassi
Université François-Rabelais de Tours
CNRS, LI EA 6300, OC ERL CNRS 6305
Tours, France
aymen.jlassi@etu.univ-tours.fr

Patrick Martineau
Université François-Rabelais de Tours
CNRS, LI EA 6300, OC ERL CNRS 6305
Tours, France
patrick.martineau@univ-tours.fr

ABSTRACT

Today, consumers request virtual resources like CPU, RAM, disk (etc.) supplied by the service providers (like Amazon) and they pay on a "pay-as-you-go" basis. Generally, the supervisors adopt virtualization technologies, which optimize resources usage and limit the operating cost. The virtualization technologies are classified in two categories. The first one concerns the heavy virtualization, which is based on virtual machines (VM) concept. Each VM emulates hardware and embeds its own operating system (OS) that is completely isolated from the host OS. The second one concerns the light virtualization, which is based on the management of containers. The containers share the host OS kernel [5] while ensuring isolation.
In this paper, we benchmark the performance and the energy consumption of an infrastructure that is based on the software Hadoop regarding the two technologies of virtualization. At first, we will identify the points to be improved concerning Hadoop performances and then we will reduce the deployment cost on the cloud. Second, the Hadoop community finds an in-depth study of the resources consumption depending on the environment of deployment. Our experiments are based on the comparison of the Docker technology (light virtualization) and VMware technology® (heavy virtualization). We come to the point that in most experiments the light technology offers better performances in completion time of workloads and it is more adapted to be used with the Hadoop software.

CCS Concepts

•Computer systems organization → Distributed architectures; •Hardware → *Platform power issues; Enterprise level and data centers power issues;*

Keywords

Virtualization, Hadoop, Resources Consumption, Benchmarks.

. .

1. INTRODUCTION

New paradigms, as BigData, Grid, Cloud computing, have been introduced to focus on powerful and parallel computation. The first means used by companies to solve the problem of the resource exploitation in the data centers is the virtualization technology, which transparently enables time-sharing and resource-sharing on servers, it aims to improve the overall productivity by letting many virtual machines running on the same physical support. In this paper, we classify virtualization tools on the heavy and light virtualization:(1)The full virtualization is based on the management of virtual machine (VM). The VM are a guest operating systems (OS) which run in parallel over physical hosts. An hypervisor ensures the interpretation of instruction from the guest OS to the host OS. (2)The light virtualization is based on the management of containers on a physical host, the containers share functions from the kernel of the host OS and have direct access to its library [5]. This classification is also used in [7]. Docker is the most sophisticated tool in its category, [13] presents the state of the art of all open source projects, which adapt Docker technology to the context of the Cloud. Traditionally, the cloud computing and big data [14] environments are mainly based on the heavy virtualization tools. The important reasons are the companies' lack of confidence in the following points: (i) the emerging technologies, (ii) the respectful efficiency of heavy virtualization (iii) the complete isolation (at different level) of the environment between guests and host OS. Nowadays, the Docker technology offers multiple capabilities of **resource isolation**. It reaches an adequate level of maturity and it can be tested with big data tools. Hadoop software is a big data environment, which was introduced by Google in 2004 [3] as parallel and distributed computation model, it is largely adopted in companies and data centers. In this work, we study and compare the **two categories of virtualization**. The experiments must be made using Docker technology, VMware technology and Hadoop software. During the evaluation, we consider *(i)* the completion time of the workloads, *(ii)* the quantity of hardware resources and *(iii)* the energy consumption criteria. We conclude then the technology that gives a cost effective cluster with a better efficiency.
The remainder is as follows. In section two, the previous studies on litterature are presented. In section three, the methodology used in the experiments is presented. In section four, the results are presented and discussed. The conclusion is presented in last section.

Table 1: Configuration of machines (physical or virtual) used in the experiments

	Host machine	Client machine
Processor	Intel Xeon(R) CPU E5-26200 @ 2.00GHz	
CPU cores	12	2 cores (4 threads)
RAM (GB)	31.5	5
HDD (GB)	500	80

2. RELATED WORKS

There are mainly two levels in the benchmark of the software Hadoop.
The first one focuses on the comparison of Hadoop with the existing engine in big data computing. For example, Pavlo et al. [12] prove that Hadoop is slower than two state-of-the-art parallel database systems, in performing a variety of analytical tasks, by a factor of 3.1 to 6.5. Zechariah et al. [4] compare Hadoop, LEMO-MR and twister (three implementations of the MapReduce model). Gu et al. [8] compare Hadoop software (HDFS/ MapReduce) to the softwares Sector/Sphere.
The second one focuses on the performances and the energetic consumption of Hadoop using different deployment architectures. For example, Kontagora et al. [11] and [16] benchmark Hadoop performances using full-virtualization (using VMware Workstation, openStack, KVM and XEN). The Docker technology is benchmarked in other contexts as the HPC technology. For example, Xavier et al. [15] present an in-depth performance evaluation of the containers based on the virtualization for HPC. They present the evaluation of the tradeoff between performance and isolation. In the same context, [6] compares the job executions using containers with executions using physical infrastructure deployment, it confirms that the overload due to the use of the container and the time completion are about 5 %. [7] compares the Hadoop software using different tools of container technology, however, neither Docker technology nor heavy technology are considered in the comparison.
Docker technology has been evaluated in the context of HPC technology, which has its specificity. In most cases, big data and HPC are two divergent fields of technologies. Each one has its own scheduling policies, resources requirements, workloads affinities. Big data infrastructure has been based on the virtualization technologies. The topic of this paper focuses on the use of Hadoop software with the Docker technology as a light virtualization tool and compares it to Vmware technology.

3. METHODOLOGY

The experiments are repeated with both types of virtualization tools. The first topic of this work is to **compare the performance variation** using the two technologies of virtualization; we compare the time completion of the used benchmarks. The second topic analyses the variation of the **energy consumption** according to the experiments and tests. We consider the time execution of the job.
In all these works, the configuration of physical host (column Host machine) and the slave machines (column Client machine) are presented in Table 1. The **slave machines** can be virtual machines or containers. To ensure the best evaluation of the platform, some configuration parameter could be fixed. For example, the rate of data replication is two and the capacity of node manager is set to 3 GB of RAM and 3 cores. The experiments are based on two levels, the first one considers two slave machines and the second one considers four slave machines. All experiments are repeated 5 times. The Ganglia software is used to recuperate monitor record LOAD, CPU, RAM metrics. It overloads the Hadoop cluster with 2 per cent [2]. The software hsflow is combined to Ganglia to retrieve I/O bound of the hard disk access on the slave machines. These metrics offer the possibility of an in-depth study in the variation in resource utilization during experiments. In order to measure energetic consumption, we use a specific engine mounted to the electrical outlet, it measures overall the energy consumption of the cluster machines every 2 seconds and save it on an external memory card.
In order to reach the topic of this work; we use the benchmarks Teragen and TeraSort and TestDFSIO. They are used by Vmware organisation; intel [9] (...) to evaluate their products. They are considered as a reference. The first kind of workloads is Teragen and TestDFSIO. They stress the hard disk and I/O resources, they are based on a set of "map" tasks which writes random data in HDFS in the a sequential manner. In these works; they generate three sizes of data 10, 15 and 20 GB using 4 slave machines. The second one is TeraSort, this benchmark stresses: memory, network and compute resources. Each data generated with Teragen is sorted with Terasort. Terasort is knwon for the capacity to aggregate output of the Teragen workload. It is based on a set of "map" tasks and "reduce" tasks. The four workloads (TestDFSIO-read, TestDFSIO-write, Teragen, Terasort) have the capacity to stress specific resource thus the evaluation results will be more accurate.

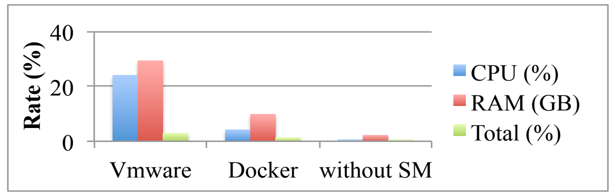

Figure 1: Resources consumption with 4 slave machines (VM or Container) without job execution

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this paper, we extract the most important results of our study. We discuss, firstly, the influence of the execution workloads on the performance and the variation in the resources i.e. CPU and I/O bounds. Secondly, we study the influence of overload on the energy consumption.

4.1 Evaluation of the Machine's Overload Capacity and CPU Bound

The overload of a machine can be defined as the difference between load of a physical machine (without any slave machine) and load after the start of slave machines on it. Figure 1 considers four slave machines (VM and container) and

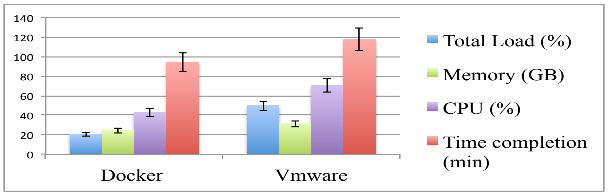

Figure 2: TeraSort execution with four slave machines and 20 GB of data

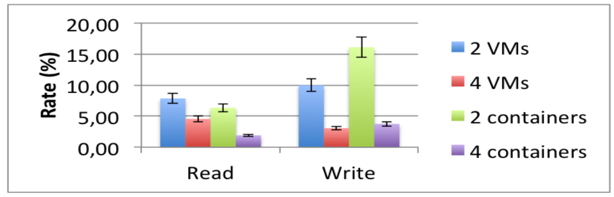

Figure 3: I/O rate with TestDFSIO (read/write) depending on the two technologies

presents: (*i*) the overload of the physical machines without the running of any slave machines (SM)(*ii*) the overload of the physical machine with 4 slaves VMs when they are idle (*iii*) the overload of the physical machine with 4 Docker containers when they are idle. We have noticed that the virtual machines reserve total configured memory since its start: thus 28 GB is booked for the cluster with 4 slaves (5 VM). But the containers use resources only when they need them, thus, 10 GB with 5 containers. This is the minimum memory needed to start hosts, guest operating systems and Hadoop daemons. It shows that the overload is measured between 3-5% for Docker containers and between 10-25% for the commercial tools. During experiments, we record overload of the physical machine. We take as an example the execution of the jobs Terasort with 20 GB of data. We remark in these conditions that container is about fourty per cent lighter than traditional virtual machine. The figure 1 illustrates the total overload and the memory consumption. Concerning the memory, the Docker technology causes less memory overload than traditional VM which reserves 100-200 MB memory per VM for hypervisor. In addition, the cluster using Docker technology is between 20 and 40% more efficient than cluster with traditional virtualization in the time completion of the jobs. We notice that using TeraSort (Figure 2); containers cause about the half of the CPU overload than traditional virtualization tool.
Based in the experiments with the two technologies, the use of two slave machines gives (between 5 and 25%) better performances than the use of four slave machines in time completion. One reason is the architecture, for example, using the cluster with two slave machines, six cores (CPU) are booked for the virtual cluster so the host OS has the six other cores to run the instructions. However, cluster with four slave machines uses 10 cores (CPU) thus only two cores are used by the host OS. We give in next part an in-depth study of Hard disk and CPU bounds exploration.

4.2 I/O-bound Variation

The TestDFSIO benchmarks are used to evaluate the HDFS health, they utilize the hard disk resource more than other resources as memory or CPU cores. The TestDFSIO benchmarks run runned on 4 slave machines, using the two technologies of virtualization. The writable data sizes vary in 10, 15 and 20 GB. The results of the job execution: TestDFSIO-write, proves that the throughput and average I/O are inversely proportional with the overload measured during the execution. For example, using 4 VMs over 20 GB of data, the overload is about 85 %, but throughput and averageI/O highly decreases. The management of hard disk bound influences directly the completion time of the workload execution. The hard disk will be critical resource and competition will also increase thus performance will decrease.
We use TestDFSIO read/write workloads to test the I/O and throughput (Figure 3). Despite the replication of data used in Hadoop (which is equal to 2), there is a difference in performance between slave machines, depending on the used technology. When we work with a cluster of four slave machines, the Docker container offers better performances in most cases in completion time. However, the difference in I/O rate is about 10% between the technologies. In the next subsection, we focus on the energetic consumption.

4.3 Slave Machines Placement Between Performance and Energetic Consumption

A research realized by the U.S. Environmental Protection Agency and the Natural Resources Defense Council [1] announced in 2007 that the cost of the energy consumption for cluster management was highest. Through the experiences, it is clear that the load and the energy consumption are proportional. When we run 4 slave machines, the overload and energy consumption increase and they are higher than the case of 2 slave machines. We can conclude that when the overload of physical machine increases (more than 85 %), the performances degrade and then, energy consumption increases. Installing many virtual machines on the physical host increases the energy consumption and they have negative influences on the performances of job execution. The overload on physical host is proportional to the number of slave machines and the workload running on them. Figure 4 shows the completion time of TeraSort workloads over a cluster with four slave machines. The cluster of two slave machines is more performant than the cluster with four machines and has a bit lower consumption than four slave machines cluster. The virtualization technology is used by the server providers to manage the load on the physical machine and to optimize energetic consumption. The overload on the physical machine is the aggregation of all overload of their guest when the VMs run in a higher load. Working with the same type of job, size of cluster and quantity of data, there is a thin difference between the use of the two virtualization tools. Docker technology consumes less energy than traditional tools. This one is caused by the use of containers instead of the overload dued to the use of virtual machines.

5. CONCLUSIONS

In this paper, we compare two virtualization tools. We deal with resource management and energy consumption on the Hadoop cluster using different types of virtualization tools.

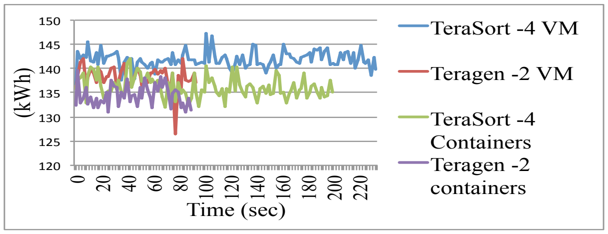

Figure 4: Energetic Consumption of different size of clusters with jobs Teragen and Terasort

A detailed study is realized to determine the influence of each resource on performance variation (refer to Section 4). The deployment of the Hadoop cluster by either using traditional virtualization or containers technology, optimizes the resource exploitation and minimizes idle resources. However, using the two technologies decreases the efficiency and the cluster performances. The main reasons are: *(i)* the overload due to the workload execution or due to the number of slave machines per host. *(ii)* The I/O scheduling, when tasks are running; a big quantity of data is transfered between the slave machines. It ensures data replication and merging between tasks. Thus, the I/O scheduling has a direct influence on the efficiency of the Hadoop cluster. The I/O environment considers the network bandwidth and harddisk access.

We consider only the resource isolation, the other kind of isolation (like user or session isolations) are not targeted in this work. The two technologies can isolate CPU, memory and Hard disk resources. In general, the container technology is equal or exceeds the traditional virtualization technology. In the major part of the test, the containers cause less overhead on CPU resources. Otherwhise, Hadoop is based on the sharing of the computing capacity between a number of slots through time. The fair share policy can increase the rate of computing resources. In addition, the containers have a performance policy to manage memory resources; free memory can be recuperated by the host operating system in order to improve general performances of the physical host. Thus, containers use specific policies to manage memory resources and computing resources, which are more adapted to the Hadoop context and offer better opportinuty to share resources than VMware.

The energy consumption is directly related to the load of the resources on the physical host i.e. the higher is the load of the physical host, the higher is the energy consumption. However, the performances depend on the number of slave machines per host and they also depend on the execution workloads.

In future, we would like to work on the optimization of the Hadoop performances by working on the scheduling policies, in order to improve performances. The approach mentioned in[10], presents the definition of the scheduling problem on the Hadoop cluster. We would also optimize the energy consumption using reports given by the Ganglia software and resource management capabilities offered by the Docker technology.

6. ACKNOWLEDGMENTS

This work was sponsored in part by the CYRES GROUP in France and French National Research Agency under the grant CIFRE n^{o} 2012/1403.

7. REFERENCES

[1] U. E. P. Agency. Report to congress on server and data center energy efficiency. 2007.

[2] I. Corporation. Intel Virtualization Technology (Intel VT). http://www.intel.com/content/www/us/en/virtualization/virtualization-technology/intel-virtualization-technology.html.

[3] J. Dean and S. Ghemawat. Mapreduce: Simplified data processing on large clusters. *Commun. ACM*, 51(1), Jan. 2008.

[4] Z. Fadika, E. Dede, M. Govindaraju, and L. Ramakrishnan. Benchmarking mapreduce implementations for application usage scenarios. In *Grid 2011*, Sept 2011.

[5] O. S. P. for Linux Containers. Getting Started with systemd. https://coreos.com/docs/launching-containers/launching/getting-started-with-systemd/.

[6] H. Gantikow, S. Klingberg, and C. Reich. Container-based virtualization for hpc. In *CLOSER 2015*, May 2015.

[7] M. Gomes Xavier, M. Veiga Neves, and C. Fonticielha de Rose. A performance comparison of container-based virtualization systems for mapreduce clusters. In *Parallel, Distributed and Network-Based Processing (PDP), 2014 22nd Euromicro International Conference on*, Feb 2014.

[8] Y. Gu and R. L. Grossman. Lessons learned from a year's worth of benchmarks of large data clouds. MTAGS '09. ACM, 2009.

[9] S. Huang, J. Huang, J. Dai, T. Xie, and B. Huang. The hibench benchmark suite: Characterization of the mapreduce-based data analysis. In *Data Engineering Workshops (ICDEW), 2010*, March 2010.

[10] A. Jlassi, P. Martineau, and V. Tkindt. Offline scheduling of map and reduce tasks on hadoop systems. In *CLOSER 2015*, May 2015.

[11] M. Kontagora and H. Gonzalez-Velez. Benchmarking a mapreduce environment on a full virtualisation platform. In *CISIS 2010*, Feb 2010.

[12] A. Pavlo, E. Paulson, A. Rasin, D. J. Abadi, D. J. DeWitt, S. Madden, and M. Stonebraker. A comparison of approaches to large-scale data analysis. SIGMOD '09, New York, NY, USA, 2009.

[13] R. Peinl and F. Holzschuher. The docker ecosystem needs consolidation. In *CLOSER 2015*, May 2015.

[14] wikipedia. Big Data definition. https://fr.wikipedia.org/wiki/Big_data/.

[15] M. Xavier, M. Neves, F. Rossi, T. Ferreto, T. Lange, and C. De Rose. Performance evaluation of container-based virtualization for high performance computing environments. In *PDP, 2013*, Feb 2013.

[16] G. Xu, F. Xu, and H. Ma. Deploying and researching hadoop in virtual machines. In *ICAL '12*, Aug 2012.